Calibrating Term Structure Models to an Initial Yield Curve

Matthew Sylvester

A dissertation submitted to the Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy.

September 16, 2020

MPhil in Mathematical Finance,
University of Cape Town.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy in the University of the Cape Town. It has not been submitted before for any degree or examination in any other University.

September 16, 2020

Abstract

The modelling of the short rate offers many advantages, with the models explored in this dissertation all offering closed-form, analytic formulae for bond prices and for options on bonds. Often, a vital primary condition is for a model to be calibrated to the initial term structure and to recover the bond prices observed in the market – that is, to be calibrated to the initial yield curve. Under the two exogenous models explored in this dissertation, the Hull-White and the CIR++, the effect of increasing the volatility parameter of the SDE increases the mean of the short rate. Increasing volatility of an SDE is a common approach to stress testing a model, as such, the consequences of bumping volatility in a calibrated model is a vital concern.

The Hull-White model and CIR++ model were calibrated to market data, with the former being able to match the observed cap prices, while the latter failed, displaying an upper bound on cap prices. Investigating this, under CIR++ model, bond option prices are shown to not be straightforward increasing functions of the volatility parameter. In fact, for high volatility, bond option prices display an upper limit before decreasing, thus providing a limit to the level of cap prices too. This dissertation points to the reason residing in the underlying CIR model from which the CIR++ is based on, and the manner in which the model is extended.

Acknowledgements

I would like to thank my mom, my dad and my sister for their love and support. Then, to AIFMRM and all those involved in the MPhil for producing the most outstanding Masters program. To my supervisor, Alex Backwell, for his time and countless insights into an ever-changing problem. And lastly to Goethe, Hesse and Jung, who have all given me so much.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Term structure modelling typically involves specifying a model through stochastic differential equations (SDEs) for various points on the instantaneous forward curve. A particular example of term structure models is short rate models, where the short rate, the shortest term instantaneous forward rate, is modelled. Through this, the rest of the term structure of interest rates is specified through the relationship between interest rates and bonds.

In specifying the dynamics of the short rate, there is an emphasis on specifying a model that reflects reality. Often of vital importance is the need for the initial term structure of a short rate model to match the one observed in the market. Intuitively, this is necessary, since bonds, being primary term structure objects should be priced correctly at the outset. Additionally, specifying dynamics of a short rate model also imply hedges overtime when the model is used to price options – thus the model should be calibrated to the instruments that will be used to hedge.

It is by focusing on these aspects that this dissertation serves to explore short rate models that are calibrated to the initial term structure. The approach taken in this introduction to follow, is to present core concepts that relate interest rates and short rate models, briefly alluded to in the preceding paragraphs, and then to focus on a specific class of models called Affine Term Structure models (ATSMs). It is within this class of models that the effect of calibrating to an initial yield curve will be explored. It is this central idea, the calibration effect, that is explored further in Chapter 2; the effect rests on the observation that increased *volatility* for models that take the initial term structure as an input, increases the *expectation* of future short rates. With this, models that possess desirable features, that are mentioned in the introduction, namely the Vasicek, Hull-White, CIR and CIR++ models, are then the focus of the rest of this dissertation. Chapter 3 serves to calibrate the models used to market data, attaining values used for simulation. Chapter 4 simulates the Vasicek and Hull-White models, comparing distributions and option prices, while Chapter 5 does similarly for the CIR and CIR++ models. Finally, Chapter 6 concludes the

discussion, synthesising the observations made in this dissertation.

1.1 Interest Rates and Term Structure Modelling

The modelling of interest rates is crucial in the pricing of not only fixed income assets such as bonds, and the associated fixed income derivatives; but also of other assets due to discounting assets over time (Benninga and Wiener, 1998).

Firstly, to understand the various interest rates involved in modelling, it may be assumed that there is a continuous term structure of zero-coupon bonds (ZCBs) with prices at t given by $P(t,T)$, where T is maturity date. Alternatively, instead of the assumption of a continuous term structure of bonds, in practice there are only observations at discrete times, producing finitely many bond prices observed in the market. These observations may then be used to imply other ZCB prices, via a bootstrapping procedure. It is this approach that will be taken in this dissertation, in order to calibrate models to market data.

The ZCBs of different tenors gives rise to different continuous interest rates which are defined as follows (Filipovic, 2009, Chapter 2):

1. continuously compounded forward rate: $R(t;T,S)$, which relates bond prices of different tenors to each other by $e^{R(t;T,S)(T-S)} = \frac{P(t,T)}{P(t,S)}$,
2. instantaneous forward rate: $f(t,T) = \lim_{S \downarrow T} R(t;T,S) = -\frac{\partial}{\partial T} \log P(t,T)$,
3. instantaneous short rate: $r_t = f(t,t)$.

Lastly, when observing bond prices, the associated yields with the bonds are also observed. These yields are inversely proportional to the price of the bond and are related by $P(t,T) = e^{-y(t,T)(T-t)}$.

The short rate is related to the bond price via risk-neutral expectation, where

$$P(t,T) = \mathbb{E}^{\mathbb{Q}}[e^{-\int_t^T r_s ds} | \mathcal{F}_t], \tag{1.1}$$

which is the time t value of a bond maturing at time T. It follows that at maturity the value of the T-Bond is $P(T,T) = 1$. To be specific, in the above equation, the time t price of the T-Bond, is equal to the expected value of 1 discounted by the bank account back to time t, under the Equivalent Martingale Measure (EMM), $\mathbb{Q}$. $\mathbb{Q}$ is also known as the Risk-Neutral Measure (RNM) since the bank account is used as the discounting asset, or the numeraire. (For the existence of $\mathbb{Q}$, and the conditions required for this expectation to hold, and a more thorough mathematical treatment, see Brigo and Mercurio (2007)).

This dissertation will centre on an approach that models the short rate, r_t, which by the expectation in 1.1 can be related to the prices of bonds, and indeed, of other payoffs with discounted expectations where the bank account is the numeraire. It is worth pointing out again something that was alluded to earlier, namely: this approach to modelling the term structure via the short rate gives access to all the different rates mentioned above. To be clear, in modelling the short rate, essentially one models one instantaneous forward rate, that of $f(t,t)$ that then this gives rise to ZCB prices, and associated yields, as well as the other forward rates of different T. This is in contrast to another class of models, the focus of which is to model different points on the instantaneous forward rate curve simultaneously. This modelling of many instantaneous forward rates is best illustrated by the Heath-Jarrow-Morton model (Filipovic, 2009).

1.2 Modelling the Short Rate

A short rate model is a model where the instantaneous short rate is specified by a SDE of the form,

$$dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dW_t,$$

often first stated under the real world measure $\mathbb{P}$. If modelling under $\mathbb{P}$, values are needed for the parameters in the drift and diffusion coefficients. It is with the aforementioned assumption that the time t price of a T-bond is given by a sufficiently smooth function F, $P(t,T) = F(t, r_t, T)$, i.e. that there is a continuous term structure of ZCBs, often implied, that one can construct a risk-less portfolio by a weighted composition of two bonds of different maturities, that a Term Structure PDE, that the function F needs to satisfy, is derived. In an arbitrage-free market, all bonds have the same market price of risk, λ. This, together with the Feynman-Kac Theorem then relates this Term Structure PDE to a SDE. That is, the price of a bond is the expectation of the following function short-rate, i.e. $P(t,T) = F(t, r(t); T) = \mathbb{E}^{\mathbb{Q}^\lambda}[e^{-\int_t^T r_s ds}|\mathcal{F}_t]$. The bond market is not necessarily complete, thus there might exist different risk-neutral measures. However, the market participants will implicitly determine a λ through bond prices. Hence, in calibrating to the observed bond prices in the market, one ensures that the SDE specified is risk-neutral. It is for this reason that in what follows, the SDEs are all specified under $\mathbb{Q}$, since calibration then ensures the correct pricing under the risk-neutral measure (Filipovic, 2009).

One of the major advantages of short rate models is that in specifying the SDE of only one explanatory variable of bond prices, r_t, there is a chance that the model may give rise to analytically tractable forms for both bond prices, and bond option

prices. For example, if the dynamics of r under $\mathbb{Q}$, gives rise to a tractable conditional distribution at time t of $e^{-\int_t^T r_s ds}$ on r_t, then bond prices may be able to be priced analytically. If one cannot deduce the exact distribution, an estimation of ZCBs numerically via some method, such as Monte Carlo Estimation can then be used. Typically, these are computationally expensive, and the history of specifying short rate dynamics has focused on the former, specifying dynamics in such a way that has led to analytically tractable formulas for bonds and other contingent claims (Brigo and Mercurio, 2007).

1.2.1 Affine Term Structure Models

After the seminal work of Vasicek (1977), which modelled the short rate with a one-dimensional source of randomness and produced analytic formulas for the pricing of ZCBs, short rate models were extended in various ways that accommodated different aspects that were desirous. A class of models was later identified by Duffie and Kan (1996), called Affine Term Structure Models, that grouped some of the models that arose after Vasicek's seminal work. The models that came to be grouped under this class of models all had the main advantage of tractability and focused on producing analytic formulas for bond prices and options (Piazzesi, 2010). This analytical tractability arose from the conditional expectation of the short rate at time t on r_t, which then simplified the conditional expectation of the bond pricing formula, Equation 1.1, to an expression only dependent on the current short rate at time t and not on the future short rates r_s, where $s \geq t$. Hence, it follows from the distributional properties of r_t that these models are defined by a relationship between bond prices and short rates that is exponentially affine, i.e., the following definition exists for these models:

Definition 1.1. A model is said to possess an Affine Term Structure (ATS) if the term structure has the form $P(t,T) = F(t, r(t); T)$, where F has the form:

$$P(t,T) = e^{A(t,T) - B(t,T) r_t},$$

where, $A(t,T)$ and $B(t,T)$ are deterministic functions.

Alternatively, it may be said that the log of the bond prices, $\log P(t,T)$, is an affine function of the short rate (Filipovic, 2009). This definition reduces the conditional expectation of $r_t, s \geq t$, seen in 1.1, to a bond price only dependent on the current value of r_t at t.

This definition is expanded upon, and is linked to the specification of the SDE of the short rate, by the following:

Proposition 1.2. *Models are ATSMs if and only if the drift and variance-squared are affine functions of the short rate. This is proved by Duffie and Kan (1996) in the paper that introduces the ATS class of models.*
That is, for short rate dynamics:

$$dr_t = \mu(t,r)dt + \sigma(t,r)dW_t,$$

both the drift, and the volatility-squared are also both affine function of the short rate, r and as such,

$$\mu(t) = \alpha(t)r + \beta(t),$$
$$\sigma^2(t) = \gamma(t)r + \delta(t).$$

This proposition makes it easier to identify and define affine short rate models, since it is easier to identify the SDE of the short rate where both the mean and volatility-squared are affine function of the short rate. It is with the use of the Term Structure PDE, and the above definition of the ATS models above, that a coupled system of differential equations are obtained; arising from the fact (1.1) relates bond prices to deterministic functions A and B. Hence,

Proposition 1.3. *assuming that μ and σ are of the affine forms mentioned above, then the model is an ATSM, where A and B satisfy the following weakly coupled ODE's:*

$$\frac{\partial}{\partial_t}B(t,T) = -1 - \alpha(t)B(t,T) + \frac{1}{2}\gamma(t)B^2(t,T),$$
$$\frac{\partial}{\partial_t}A(t,T) = \beta(t)B(t,T) - \frac{1}{2}\delta(t)B^2(t,T),$$

with initial conditions, $A(T,T) = B(T,T) = 0$ (Björk, 2009).

When the SDE of r is in terms of $\alpha, \beta, \gamma, \delta$, the solutions to the ODEs give $A(t,T)$ and $B(t,T)$. However, the solving of these equations is not always trivial, but provides a key method for finding analytic expressions for bond prices. Another method, as mentioned before, is the distributional properties of r_t that are of interest in the pricing of bonds. An example of this that will be seen later is: when the instantaneous short rate is Gaussian, it is rather straightforward to calculate the expected value of a log-normal stochastic variable via the Normal moment-generating-function (MGF). In short, ATSM have the added advantage of having the above system of ODEs, which can be seen to be easier to solve to obtain bond prices, and to price derivatives via $dP(t,T)$ and change of measure techniques (Björk, 2009).

1.3 The Chronology of Single Factor Affine Models

Vasicek (1977) in his seminal paper assumed that the instantaneous spot rate evolved as an Ornstein-Uhlenbeck process under the real world measure $\mathbb{P}$, with constant coefficients.

$$dr_t = k(\theta - r_t)dt + \sigma dW_t. \tag{1.2}$$

This SDE for the short rate accounted for the useful empirical observation of mean reversion, where θ is the long-term value for the short rate, or the mean-reverting rate and is used to incorporate a view, or an anticipation of future events. The distribution of the short rate is Gaussian, which has the benefit of analytic tractability for bond prices and options on bonds. The rate of mean reversion is k; intuitively, regarding the drift term, if r_t is below θ, the drift term would be positive, pulling the evolution of the process up towards the mean-reversion level θ by the speed of reversion, k. If r_t is 0 at some time t, the SDE allows the short rate to still rise, or fall in the future, moving away from 0. Analogously, the interpretation for $r_t \geq \theta$ follows. The possibility of negative instantaneous short rates resulting from the normal distribution has been seen as a drawback in the development of short rate models, as it results in can result in an error in pricing particular instruments.

Cox *et al.* (1985) introduced the square root term in the diffusion coefficient, overcoming precisely this possibility of negative instantaneous short rates in the Vasicek model. Addressing that, and still maintaining the property of θ anticipating future events, the CIR model models the short rate with the following dynamics:

$$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t. \tag{1.3}$$

The SDE above removed the possibility of negative rates by the introduction of the square root of r_t in the volatility component, provided certain conditions on the other parameters are met. The Feller condition, $2k\theta \geq \sigma^2$, ensures that the origin is inaccessible, thus the process remains positive since the upward drift will then outweigh the contribution to the change in r_t from the diffusion coefficient – necessary because of the square root in the above SDE.The parameters k and θ have a similar interpretation as in the Vasicek model, however of particular importance, is the effect of the square root of the short rate in the diffusion term; it is this that results in the variance of interest rates increasing with increasing rates Cox *et al.* (1985).

As mentioned before, the strength of these early models rests on tractability and in pricing of both bonds and options on bonds analytically. The distribution of r_t in the CIR model is non-central Chi-squared. The parameter values for both models

are decided both by what best matches the initial term structure seen in the market, inverted from the observed bond prices. It may be argued then, that the empirical evidence that might be used to specify a particular value – say previous data indicates a high volatility of rates, leading to a high choice of σ, or by using past market data, a particular mean-reverting parameter θ is apparent, are somewhat offset depending on the use of the model. As mentioned earlier in the necessity to calibrate to the initial term structure in order to correctly specify an SDE under risk-neutral dynamics, done implicitly through calibration of parameter values. Consequently, in the case of the above two models, the unknown parameter values are solved for by minimising (absolute) difference between the bond prices produced by the model and ZCBs observed in the market. A minimsation scheme could also be stipulated on the associated yields, leading to slightly different results.

The two models introduced, the Vasicek and the CIR, are characterised as endogenous term-structure models (Brigo and Mercurio, 2007), which means that the current term structure of interest rates is computed by the specification of these models and given as an output – it is considered as endogenous to the model, rather than as an input. Typically, it is a necessary concern that the initial term structure of continuously compounded rates that the model produces matches the initial term structure observed in the market as closely as possible. However, depending on how the parameters are chosen, typically with this idea of θ reflecting some anticipation of future events, σ and k might not be able to match the initial term structure, typically minimising the difference between the bond prices implied by the model and that observed in the market, despite the efforts made to minimise the absolute difference between two. If the initial term structure of the model does not match that of the observed market's, all pricing and use of the model for hedging is immediately incorrect. Thus, a natural consequence of this is the demand that models can fit the observed term structure completely, for pricing and hedging purposes in particular (Björk, 2009).

Exogenous models arose to meet this need, and it is in these models where the initial term structure is taken as an input into the model – that is, the initial term structure is exogenous to the model, hence the name. This is done by introducing parameters that are of infinite dimension, instead of the scalars seen above. Thus, by varying the parameter deterministically in the SDE, i.e., by letting it be time dependent, the required system of equations may be solved (Björk, 2009). The models that are exogenous and possess an ATS follow below.

Ho and Lee (1986) were the first to introduce an exogenous term structure model, with the Ho-Lee model having the following dynamics,

$$dr_t = \theta(t)dt + \sigma dW_t.$$

However, the lack of a mean reversion term in the above SDE limited its use, as it lacked the same empirical justification that the Vasicek and CIR models rested on. As such, Hull and White (1990) extended the Vasicek model by introducing time varying parameters, giving rise to the Hull-White model, sometimes known as the Extended Vasicek model. Following from the explanation outlined in Brigo and Mercurio (2007), focusing on the deterministic extension of only one of the parameters, instead of time-varying volatility and mean reversion too, the dynamics of the short rate are specified as follows:

$$dr_t = (b(t) - ar_t)dt + \sigma dW_t, \tag{1.4}$$

where, a and σ are constants and are chosen to fit a viable volatility structure, and b is a deterministic function of time, chosen such that it fits the observed initial term structure of bonds $P(0,T)$ (Brigo and Mercurio, 2007).

Placing the Hull-White parametrisation in an equivalent form, to compare to the Vasicek model, manipulation is as follows:

$$dr_t = k(\theta(t) - r_t)dt + \sigma dW_t, \tag{1.5}$$

where $\theta(t) = \frac{b(t)}{k}$, and $k = a$, a simple substituting to move from the definition found in Brigo and Mercurio (2007) to a comparable form with Vasicek and models presented later – the last substitution of k being an aesthetic one.

It is in this model that an ideal observation of the calibration effect might be most easily identifiable, and it is this effect that this dissertation will seek to explore. To be more specific with regards to this effect, it is by fitting the initial term structure with the deterministic parameter specified in the Hull-White model, that the two other parameters may then be adjusted, effecting the SDE dynamics. This dissertation then focuses on the effect on short rates when precisely these parameters are changed, in particular when bumping σ in combination with these changes – the effect of calibrating is thus teased out. It is important to note that the parameter values that allow this SDE to be specified under $\mathbb{Q}$ are motivated precisely by calibrating to market data of a particular volatility structure. It is by calibrating to an initial term structure, and manipulating parameter values for the other 'constant' parameters that the next chapter explores the effect of.

Lastly, how the Vasicek model was extended to the Hull-White model, moving from endogenous to exogenous, the CIR model may be extended too. If done similarly, then the CIR model would ideally make the θ parameter in Equation 1.3 time varying. However, as Brigo and Mercurio (2000) explains, this does not give rise to analytically tractable formulae. Nor does the scenario where all the coefficients – both $k(t)$, and $\sigma(t)$ too – are time varying. Numerical solutions are needed for

these equations to be solved, and Brigo and Mercurio (2000) then reasons that this model loses the analytical tractability that placed it above log-normal models. As an alternative, Brigo and Mercurio (2000) introduces the CIR++ model, which creates an endogenous extension of the CIR model by means of a deterministic shift. The SDE is specified as:

$$dx_t = k(\theta - x_t)dt + \sigma\sqrt{x_t}dW_t, \tag{1.6}$$

$$r_t = x_t + \varphi^{\text{CIR}}(t), \tag{1.7}$$

where the deterministic shift function is given by

$$\varphi^{\text{CIR}}(t) = f^M(0,t) - f^{\text{CIR}}(0,t),$$

$$f^{\text{CIR}}(0,t) = \frac{2k\theta(e^{th}-1)}{2h+(k+h)(e^{th}-1)} + x_0\frac{4h^2e^{th}}{(2h+(k+h)(e^{th}-1))^2},$$

where $h = \sqrt{k^2+2\sigma^2}$, and $f^M(0,t)$ is the market observed instantaneous forward curve.

Note that in this model the process x_t is exactly the same as the process specified in the CIR SDE (1.3). The CIR++ can be understood intuitively as shift in this random process, such that the model prices the observed ZCBs in the market from the outset. For a full discussion on the CIR++ model, and how other short rate models may be extended in a similar manner, Brigo and Mercurio (2000) and Brigo and Mercurio (2007) provide detailed descriptions of the process of doing just this.

Chapter 2

On the Calibration Effect, Short Rate Models

The primary aim of this dissertation, as stated in the introduction, is to explore the calibration of term structure models, and the effects of this calibration. As can be seen in the above literature review, the Hull-White model is a key starting point when deducing the effect of the calibration effect in the presence of constant volatility. The idea behind exploring single-factor models, is to explore the effect of introducing a time deterministic parameter that takes the initial term structure as an input. It is this that changes the meaning behind the parameters used in the SDE; no longer is there a mean reversion rate θ, but rather some $b(t)$ that enables exact calibration to an initial term structure - no longer is it 'mean reverting' to a rate, but rather to the time-determistic parameter dependent on the initial yield curve. As shall be seen, it is with the resulting effect of calibration on the SDE and the future evolution of the short rate that the effect of changing the other parameters in the SDE is explored.

As a base case, considering the Vasicek model, and then proceeding in extending this model to the Hull-White model, an illustration emerges of this effect.

2.1 Motivation for the Calibration Effect

The SDE of the Vasicek Model is defined in 1.2, from which the derivation of conditional expectation and variance of the short rate at some time t is:

$$\mathbb{E}[r_t|\mathcal{F}_s] = r_s e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}), \tag{2.1}$$

$$\mathbb{V}[r_t|\mathcal{F}_s] = \frac{\sigma^2}{2k}(1 - e^{-2k(t-s)}). \tag{2.2}$$

It can be seen that this expectation of the future rate is not dependent on the choice of the σ parameter, the volatility specified in the SDE.

Unconditionally, this simplifies to

$$\mathbb{E}[r_t] = r_0 e^{-kt} + \theta(1 - e^{-kt}),$$
$$\mathbb{V}[r_t] = \frac{\sigma^2}{2k}(1 - e^{-2kt}).$$

The SDE of the Hull-White model is defined as in 1.4, from which the derivation of the conditional distribution, following Brigo and Mercurio (2007), at some future time t has mean and variance

$$\mathbb{E}[r_t|\mathcal{F}_s] = r_s e^{-k(t-s)} + \alpha(t) - \alpha(s)e^{-k(t-s)}, \tag{2.3}$$
$$\mathbb{V}[r_t|\mathcal{F}_s] = \frac{\sigma^2}{2k}(1 - e^{-2k(t-s)}), \tag{2.4}$$

where $\alpha(t) = f^M(0,t) + \frac{\sigma^2}{2k^2}(1 - e^{-kt})$, and $f^M(0,t)$ is the observed instantaneous forward rate curve that is implied at time 0. This forward rate curve is, in practice, constructed from bootstrapping the forward curve from observed bond prices, which will be done in the following chapter.

Unconditionally, this simplifies to

$$\mathbb{E}[r_t] = f^M(0,t) + \frac{\sigma^2}{2k^2}(1 - e^{-kt})^2,$$
$$\mathbb{V}[r_t] = \frac{\sigma^2}{2k}(1 - e^{-2kt}),$$

from which it is slightly easier to observe the effects of the changing σ parameter on the expectation of r_t.

Comparing the SDEs and the resulting unconditional means at some future time of the short rate, the effect of changing the parameters of the SDE on the distributional properties of the short rate are deduced. In the case of the Vasicek model, any change in the model parameter σ has no effect on the mean of the short rate. There is however an increase in the variance; an increase in the value of σ simply increases the paths that lie on either side of the mean – that is, when generating a sample of the short rate at t, there will be a greater spread of paths located around this unchanged mean. Comparatively, in the case of the Hull-White Model, the mean at some future time of the short rate, is now dependent on the volatility parameter σ. Here, an increase in the value of σ, increases the value of the mean - that is, we can think of an increase in σ as increasing the 'trajectory' of more of these future paths. It is this observation that is of particular relevance when considering the effects of changing σ on the values of the corresponding theoretical options on the short rate, and on corresponding options on bond prices.

The theoretical considerations serve as a motivation for the exploring the effects of calibration: by extending Vasicek model to allow for the inclusion of the initial

yield curve as an input parameter, one makes the the conditional expectation of the short rate dependent on σ, rendering the future expectation of the short rate dependent on the volatility specified in the SDE.

It is with this that one can consider the effect by creating a 'theoretical' options on the short rate. The term theoretical is used, since this is not an option that is traded in the market, as the payoff would be based on an instantaneous short rate realisation, a theoretical interest rate in itself. In any case, consider a:

- Call Option on r_t, with payoff of $(r_T - K)^+$ at T,
- Put Option on r_t, with payoff of $(K - r_T)^+$ at T,

where $(\cdot)^+ = \max(\cdot, 0)$.

In constructing these options on the observed instantaneous short rate at the future time T, the effects of changing σ can be deduced. In the case of a call option on r_T, more of the paths will be pulled into the money at the same strike with increasing σ, since σ will now increase both the mean and the spread of the short rates about that mean, and so the price of the call will increase. Interestingly, with a put, more of the paths will be pulled out of the money with the increasing mean as a result of the increasing σ, and so the price of the put will be expected to decrease due to this effect. However, increasing this variance, might also have the effect of increasing the spread about this new mean, and place more paths in the money for the put too, which would counteract the increasing mean (and for the call, that increased spread might offset the increase in mean). Hence, in the case of the Hull-White model, the effect of calibrating to an initial term structure, has a crucial first consequence – that call options on the instantaneous short rate might increase with increasing σ since both variance and mean increase, while put options might decrease in value with increasing σ depending on the interaction between increased mean and variance.

This primary comparison with a theoretical short rate option is indeed of intuitive importance, but importantly, needs to be aligned with the pricing of an option of some practical value. It is with this need that the effect of increasing σ on bond options is also made. As mentioned before, the ATS class of models, has the advantage of analytical tractability, and as such, there are closed form solutions for options on bonds. This then serves as the motivation to explore the effects of changing the parameter in the Hull-White model. The point of illustrating the intuition using the Vasicek model first, served two purposes. Firstly, it was to establish the heuristic, or the intuitive assumption that we have from most models – namely, that increasing σ increases option prices. Secondly, it was to draw attention to how making one parameter deterministic brings one to the Hull-White SDE, and how

the effect of changing σ on the short rate is not immediately obvious. The purpose of the instantaneous short rate option is not necessarily to price the option itself, but rather through the interpretation of the payoff of these options, to understand the distributional effects of moving from an endogenous to an exogenous model.

With that brief exposition of the calibration effect on the distributional properties of r_t, an exposition for the CIR and CIR++ models should follow in an analogous manner as to the Vasicek, Hull-White case. That being said, the non-central chi-squared distribution of the CIR short rate, following from the SDE defined in 1.3, gives the following conditional mean and variance:

$$\begin{aligned}\mathbb{E}[r_t|\mathcal{F}_s] &= r_s e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}),\\ \mathbb{V}[r_t|\mathcal{F}_s] &= r_s \frac{\sigma^2}{k}(e^{-k(t-s)} - e^{-2k(t-s)}) + \theta\frac{\sigma^2}{2k}(1 - e^{-k(t-s)})^2,\end{aligned}$$

where $t \geq s$ (Cox *et al.*, 1985).

Expressing this unconditionally, again provides an easier and more useful comparison

$$\begin{aligned}\mathbb{E}[r_t] &= r_0 e^{-kt} + \theta(1 - e^{-kt}),\\ \mathbb{V}[r_t] &= r_0 \frac{\sigma^2}{k}(e^{-kt} - e^{-2kt}) + \theta\frac{\sigma^2}{2k}(1 - e^{-kt})^2.\end{aligned}$$

The appropriate comparison with the CIR, is that with the other endogenous model, the Vasicek model. Firstly, in comparing the expectations of the short rate, one observes that they are the same under both models. As such, when $k \to \infty$, the expectation of r_t tends to θ, as one would expect; the higher mean reversion rate overcoming the effects of the diffusion term. When $k \to 0$, $\mathbb{E}[r_t|r_s] \to r_s$, the current interest rate, as a value of $k = 0$ would ensure that the drift term contributes nothing.

Secondly, in comparing the variances of the short rate, these expressions are not equal. However, considering the same case when $k \to \infty$, under both models, $\mathbb{V}[r_t] \to 0$. Comparatively, when $k \to 0$, a slightly closer inspection is required using l'Hôpital's rule to evaluate these limits of indeterminate form.

For the Vasicek model,

$$\begin{aligned}\lim_{k\to 0} \mathbb{V}[r_t|\mathcal{F}_s] &= \lim_{k\to 0} \frac{\sigma^2}{2k}(1 - e^{-2k(t-s)})\\ &= \frac{\sigma^2}{2} \lim_{k\to 0} \frac{-e^{-2k(t-s)}(-2(t-s))}{1} \text{ (by l'Hôpital's)}\\ &= \sigma^2(t-s).\end{aligned}$$

For the CIR model,

$$\begin{aligned}\lim_{k\to 0}\mathbb{V}[r_t|\mathcal{F}_s] &= \lim_{k\to 0}\left(r_s\frac{\sigma^2}{k}(e^{-k(t-s)}-e^{-2k(t-s)})+\theta\frac{\sigma^2}{2k}(1-e^{-k(t-s)})^2\right)\\ &= \lim_{k\to 0}\left(r_s\sigma^2(-(t-s))\frac{(e^{-k(t-s)}-2e^{-2k(t-s)})}{1}\right)+\\ &\lim_{k\to 0}\left(\theta\sigma^2\frac{2(1-e^{-k(t-s)})(t-s)e^{-k(t-s)}}{2}\right)\\ &= r_s\sigma^2(t-s),\end{aligned}$$

where the second equality followed by applying l'Hôpital's rule, and the last line by applying the limit. Comparing both expression, the presence of r_s, the current short rate, in the variance of the CIR follows from the diffusion coefficient in the SDE.

While comparison by placing parameter values in both expressions might be a more thorough way of comparing the two, some further intuition can be gleaned from a similar reasoning mentioned above, regarding the presence of the square root of r_t in the diffusion coefficient in 1.3: since r_t is an instantaneous short rate, and is typically less than 1, the square root of it, while making it slightly larger, will still decrease the variance of the path generated for an equivalent value of σ, if that same σ is used in the Vasicek model. This is clearly seen in the comparison between the variances above, when $k \to \infty$, all else equal that is, one would expect the short rate at time t to have a lower variance in the CIR model, than the Vasicek model.

The CIR++ model, as mentioned before, is a deterministic shifted CIR model. The mean of the r_t under the of the CIR++ model is then simply that of the the CIR model, shifted upwards by the function φ^{CIR}, since this function is increasing in terms of σ, as simple substitution will reveal in 1.7.

Chapter 3

Bootstrapping, Calibration and Simulation

The short rate models that were introduced in the brief literature review, and followed by observations of their distributional properties, are now explored with changing σ. Alluded to before, an intuition was established to indicate that the 'value' of σ is not the same under both models. Calibration then to the market can be seen in this dissertation as an attempt to find the analogous parameters for each model, in order to observe the effect of changing σ.

Calibration to the market, is also an attempt to 'recover' an equivalent martingale measure, in particular, the risk-neutral measure. As mentioned before, this measure is not necessarily unique and by calibrating to the market, recovery of the one the market participants collectively imply through price making is possible, allowing the SDEs to be specified under $\mathbb{Q}$. The aim of this section is to calibrate model parameters k and σ for the Hull-White and CIR++. The parameter values found for the Hull-White and for the CIR++ model, are then passed down to the Vasicek and CIR models respectively. The reason for this, and also for not calibrating the other parameters, θ and x_0, is that the former is an anticipation of future events, and an observation of current interest rate r_0 for the latter. As such, one would set θ to be the longest term yield observed from bootstrapped market data, and set x_0 to r_0.

3.1 Bootstrapping the Initial Yield Curve

Both the Hull-White model, and the CIR++ model, take the initial yield curve as an input, thus both need the observed instantaneous forward rate as an initial input into the model before calibration is possible. The observed instantaneous forward rate needs to be taken from observed market data. The market data in question here, is that of Euro 6-month Curve Structure on the 29 July, 2015. Here, the 6-

month Euribor is the shortest instrument quoted, followed by the 6x12 FRA, 12x18 FRA, and then semi-annual swaps from a maturity of 2 out to 10 years.

The bootstrapping procedure, of which there are many alternative approaches, followed in this dissertation was to simply recover the discount factor associated with the desposit rate and the FRAs, and then to assume a constant multiplicative factor between the implied ZCBs for each of the maturities after 2 years – this means, when considering the discount factors, or ZCBs, at $t = 2$, $Z_{2.5} = x_2 \times Z_2$, and $Z_3 = x_2 \times Z_{2.5}$, or equivalently, $Z_3 = {x_2}^2 \times Z_2$.

Once the discount factors are calculated in this sequential procedure, bootstrapped that is, from the curve structure, then continuously compounded and instantaneous forward rates are calculated.

A brief description of the instruments used in bootstrapping is useful. Firstly, the shortest term deposit rate, is simply a quote of the simple rate at which one may deposit money today, earning the associated interest in 6 months' time. Secondly, a Forward-Rate Agreement (FRA), agrees on a fixed interest rate for a specific period in the future. In the case of the first FRA, a 6x12, this means that the rate the market trades at today is an agreement for that simple rate of interest to apply between 6 and 12 months – what this contract in essence allows, is for the holder to trade an unknown variable interest rate that will only become know at some future time, for a fixed interest rate that can be 'locked-in', or fixed, today. Finally, an Interest Rate Swap, or Swap, is a generalisation of the FRA, most intuitively thought of as a sum of adjacent FRAs traded at one particular rate. Here, there are reset days, at which fixed and variable interest rate payments are exchanged. For a full treatment of the relevant pricing formula, see Chapter 1 of Brigo and Mercurio (2007), which features a more thorough yet still concise introduction.

The market data consists of the traded instrument and the associated market rates at which they trade. The discount factors are then calculated, and interpolated on, giving rise to the following bootstrapped data set:

t	Instrument	Market Rate	$Z(0,t)$	x_i	$r(0,t)$	$f(0,t)$
0	-	-	1.00000	-	-	0.000480
0.5	6-month Euribor	0.0004800	0.99976	-	0.000480	0.000480
1	6x12 FRA	0.0006000	0.99946	-	0.000540	0.000600
1.5	12x18 FRA	0.0010500	0.99894	-	0.000710	0.001050
2	2yr Swap	0.0009850	0.99803	-	0.000985	0.001811
2.5		-	0.99640	0.99836	0.001445	0.003283
3	3yr Swap	0.0017500	0.99476	-	0.001751	0.003283
3.5		-	0.99166	0.996884	0.002393	0.006242
4	4yr Swap	0.0028689	0.98857	-	0.002874	0.006242
4.5		-	0.98385	0.995228	0.003617	0.009567
5	5yr Swap	0.0041974	0.97916	-	0.004212	0.009567
5.5		-	0.97301	0.993717	0.004975	0.012605
6	6yr Swap	0.0055780	0.96689	-	0.005611	0.012605
6.5		-	0.95944	0.992295	0.006370	0.015471
7	7yr Swap	0.0069583	0.95205	-	0.007020	0.015471
7.5		-	0.94359	0.991113	0.007742	0.017854
8	8yr Swap	0.0082750	0.93520	-	0.008374	0.017854
8.5		-	0.92611	0.990275	0.009031	0.019545
9	9yr Swap	0.0094719	0.91710	-	0.009615	0.019545
9.5		-	0.90782	0.989876	0.010180	0.020351
10	10yr Swap	0.0104993	0.89863	-	0.010689	0.020351

Tab. 3.1: Bootstrapped initial yield curve – Euro 6-month curve structure

where, $f(0,0) = f(0,0.5)$. That is, the initial short rate $r_0 = f(0,0)$ is set to equal the yield on the shortest instrument, the 6-month deposit. Furthermore, $\theta = 0.010499253$, the longest term rate.

Rates, or yields, can be calculated from the discount factor by the one-to-one relationship between bond prices and yields leading to

$$Z(0,t) = e^{-r(0,t)(t-0)}.$$

Instantaneous forward rates are calculated from the discount factors by relating

each successive discount factor to each other as follows,

$$\begin{aligned} Z(0,t_{i+1}) &= Z(0,t_i)e^{-\int_{t_i}^{t_{i+1}} f(t_i,t)dt} \\ &= Z(0,t_i)e^{-f(t_i)(t_{i+1}-t_i)}, \end{aligned}$$

where the last line follows by assuming an equal continuous instantaneous forward rate seen from $t = 0$, between each interval $[t_i, t_{i+1}], \forall i$.

In plotting the ZCB prices, and the implied initial instantaneous forward rates, the following are observed:

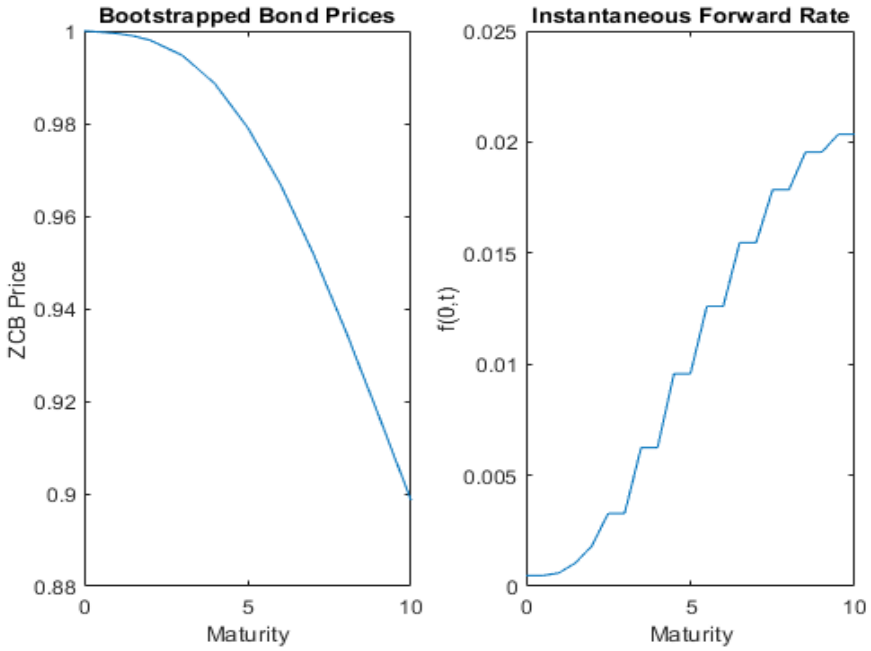

Fig. 3.1: Bootstrapped Rates from Market Data.

3.2 Calibration to the Initial Yield Curve and Volatility Structure

The cap market prices provide information for a volatility structure, to which calibration is done. This is to find k, σ for both the Hull-White and CIR++ models. The market quotes these cap prices in the form of an implied vol, which is the vol parameter that will relate Black's price to the price of the cap quoted in the market – it is market practice to quote the price of a cap via the implied vol.

In calculating a cap price, Black's formula is applied to the individual caplets that make up the cap. Again, Brigo and Mercurio (2007) can be used to find Black's

formula for calculating the price of a cap from the implied volatility. Following this, the cap prices are quoted in the table below.

Cap Maturity	Implied Vol	Cap Price
2	610.35	0.001715076694630
3	306.43	0.004673483733910
4	154	0.008733363580603
5	118.96	0.015115729622176
6	98.33	0.022972753686040
7	84.64	0.032034205419105
8	77.92	0.042799203558884
9	72.51	0.054085877626734
10	67.42	0.064935013163585

Tab. 3.2: Market cap implied vol, and implied cap price

It is with the cap prices that the parameters for k, σ can be found that minimises the absolute difference between the market cap prices and the relevant model cap prices for the term structure of caps. Formula for pricing caps under the Hull-White and CIR++ models can be found in Brigo and Mercurio (2007) Chapter 3. Following the minimisation, the following parameters for each model where found to best match the market cap prices:

	k	σ
Hull-White	0.009570405184446	0.006656075284058
*CIR++	0.396184075857206	0.358734777471847
CIR++	0.763183338079797	0.124396038356878

Tab. 3.3: Calibration procedure: parameter values

where, a *CIR++ refers to a calibration where the Feller condition, $2k\theta \geq \sigma^2$, was relaxed, and CIR++ was where this condition was enforced. The sum of the absolute difference between the cap prices produced by the models, and that of the market are found in the table below:

	Sum of Abs Difference
Hull-White	0.028683909213524
*CIR++	0.041465996629685
CIR++	0.075711780085592

As can be seen, in the table above and more explicitly in the table to follow, the

Hull-White model is roughly able to match the cap prices observed in the market and performs far better than the CIR++ model. Naturally, finding two parameters while trying to match 9 cap prices will rarely provide a perfect fit across the entire term structure.

Cap Maturity	Market Cap Price	Hull-White Cap Price
2	0.00171507669463000	0.00399810663098275
3	0.00467348373391000	0.00817364855408802
4	0.00873336358060300	0.0133996044077104
5	0.0151157296221760	0.0196228813088582
6	0.0229727536860400	0.0266904145542957
7	0.0320342054191050	0.0344822121008302
8	0.0427992035588840	0.0427991872539086
9	0.0540858776267340	0.0514047226677948
10	0.0649350131635850	0.0600545300342768

Tab. 3.4: Calibration procedure: Hull-White cap prices

For the CIR++ model, enforcing the Feller condition, results in cap prices that simply do not match that of the market – calibration has failed in this case. In relaxing the Feller condition, here too, the CIR++ model seems unable to produce cap prices as high as the market – this issue will be explored in Chapter 5.

Cap Maturity	Market Cap Price	*CIR++ Cap Price	CIR++ Cap Price
2	0.00171507669463000	0.00396809994771242	0.00270374264266211
3	0.00467348373391000	0.00835040994297384	0.00560280936300939
4	0.00873336358060300	0.0134774542393617	0.00914268080163767
5	0.0151157296221760	0.0190640223843363	0.0133476594864128
6	0.0229727536860400	0.0250732315963645	0.0181306361203770
7	0.0320342054191050	0.0320342053957091	0.0234687940134150
8	0.0427992035588840	0.0389274522154714	0.0291018671115293
9	0.0540858776267340	0.0459473460915975	0.0346751979840594
10	0.0649350131635850	0.0522021102292349	0.0398341570733051

Tab. 3.5: Calibration procedure: CIR++ cap prices

3.3 Simulation Specifics

3.3.1 Model Simulation

In the case of the Vasicek model, the SDE (1.2) may be integrated to give the expression,

$$r_t = r_s e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}) + \sigma \int_s^t e^{-k(t-u)} dW_u,$$

where $s \leq t$ (Brigo and Mercurio, 2007). The last term can be deduced to be normally distributed, as the integrand is deterministic. The integral is thus normally distributed with mean zero via the martingale property, and the variance is deduced via Ito isometry. As such, a discretisation may then be obtained replacing the last expression with a normal random variable. Hence, letting $\Delta t = t_{i+1} - t_i$, this gives

$$r_{t_{i+1}} = r_{t_i} e^{-k\Delta t} + \theta(1 - e^{-k\Delta t}) + \sqrt{\frac{\sigma^2}{2k}(1 - e^{-2k\Delta t})} Z_i,$$

where $Z \sim N(0, 1)$. In this expression, the first two terms are the conditional mean, seen in (2.1), added to the random component of the last term that generates the required conditional variance, seen in (2.2). This algorithm is an exact simulation of the stochastic process, i.e., it produces precisely the distribution of the Vasicek process at each interval time t_i. Alternatively, a slightly simpler Euler scheme might be used to generate the Vasicek short rate. This Euler discretisation is expressed analogously as

$$r_{t_{i+1}} = r_{t_i} + k(\theta - r_{t_i})\Delta t + \sigma\sqrt{\Delta t} Z_i,$$

which has some discretisation error, but since the expression itself is normal, the distribution of the short rate should closely follow the true distribution even over wider intervals. This Euler discretisation has been used to provide another validation step in the implementation of code in this dissertation (Glasserman, 2013).

In the case of the Hull-White model, an expression of r_t dependent on r_s is:

$$r_t = r_s e^{-a(t-s)} + \alpha(t) - \alpha(s)(1 - e^{-a(t-s)})^2 + \sigma \int_s^t e^{-a(t-u)} dW_u,$$

where $\alpha(t)$ is defined as in Equation 2.3. As such, paths of the instantaneous short rate may be generated by relating the short rate at some time t_i to t_{i-1}:

$$r_{t_i} = r_{t_{i-1}} e^{-a\Delta t} + \alpha(t_i) - \alpha(t_{i-1})(1 - e^{-a\Delta t})^2 + \sigma \int_{t_{i-1}}^{t_i} e^{-a(t-u)} dW_u.$$

Proceeding in the same manner as above, replacing the Ito integral with a normal random variable, the following expression is used in the code to simulate sample

paths of the Hull-White short rate:

$$r_{t_i} = r_{t_{i-1}} e^{-a\Delta t} + \alpha(t_i) - \alpha(t_{i-1})(1 - e^{-a\Delta t})^2 + \sqrt{\frac{\sigma^2}{2a}(1 - e^{-2a\Delta t})} Z_i,$$

where $Z \sim N(0, 1)$. An Euler expression may be similarly derived as in the case of the Vasicek, but the need to deterministically define $b(t)$ at each t makes this Euler expression slightly more complicated to use than the above expression. Since the former is exact and easier, it is favoured.

For the CIR Model, an Euler discretisation is used. Since the Euler discretisation is normal conditional on the previous of short rate, it only approximates the short rate, which, as mentioned, has a non-central chi-squared distribution. An Euler discretisation of the CIR process is complicated by the presence of the square root of the short rate in the SDE (1.3). If the process r_t goes negative at some point in the simulation, then the square root will be non-real. In order to avoid taking the square root of a negative number, there are a handful of schemes that may be used to discretise the CIR process. Lord *et al.* (2010) discuss these various methods of discretely simulating a CIR process, doing so in the context of stochastic volatility models. Lord *et al.* (2010) show that the 'full truncation' method has the best convergence properties among the various possible alternatives.

Applying the full truncation to the CIR process gives the following discretisation scheme:

$$\begin{aligned} r_{\text{temp}} &= r_{t_i} + k(\theta - r_{t_i})\Delta t + \sigma\sqrt{r_{t_i}}\sqrt{\Delta t} Z_i. \\ r_{t_{i+1}} &= \max(r_{\text{temp}}, 0). \end{aligned}$$

A discussion of the Feller condition, how the different schemes approximate the CIR process, and the effects on pricing is not within the scope of this dissertation, and relevant information is perhaps best found first in Cox *et al.* (1985) and Brigo and Mercurio (2000), and Lord *et al.* (2010).

The CIR++ model follows an identical simulation as that of the CIR process. This can be seen in the SDE, (1.7) where x_t in the CIR++ is the same stochastic process as r_t in the CIR model. It is the function φ that shifts the CIR paths to that of the CIR++, in order to match the initial yield curve. A point worth returning to, is that an Euler scheme used to generate the CIR stochastic process, not only has a discretisation error which is also found in an Euler discretisation of the Vasicek and Hull-White processes, but also that of a 'mismatch' between distributions. As such, the discretisation time steps need to be finer, or fine enough at least, to approximate the true distribution. Hence, in discretising the CIR process, and similarly the CIR++ process, a time step of $\frac{\Delta T}{10}$ is used. That is, if the time step of the normally distributed models of Vasicek and Hull-White were that of the 6-month time-step

given by the market data, i.e, $\Delta t = 0.5$, half a year, then the CIR discretisation subdivides each year into 20 time steps.

3.3.2 Random Variable Calculation

The consideration of theoretical options briefly alluded to earlier, that of these instantaneous short rate options, have the following payoffs:

- Call option on r_t, with payoff of $(r_T - K)^+$ at T,
- Put option on r_t, with payoff of $(K - r_T)^+$ at T.

Thus, in order to calculate a price at $t = 0$, a discounted expectation is needed. Since the bank account which acts as the numeraire under risk-neutral expectation, is itself a stochastic process, both the payoff and the discounting numeraire depend on the SDE specified. Thus in pricing this option, the discounted payoffs are:

- $c_0 = \mathbb{E}^{\mathbb{Q}}[e^{-\int_0^T r_s ds}(r_T - K)^+]$,
- $p_0 = \mathbb{E}^{\mathbb{Q}}[e^{-\int_0^T r_s ds}(K - r_T)^+]$.

In using a Monte-Carlo estimation, there is then a need to calculate $Z_T = e^{-\int_0^T r_s ds}$ for each path at some time T. Observing that in calibrating a model, and moving from an endogenous to an exogenous model, there is an increase in the mean of the short rates, there will be a resulting effect on the discount factor Z_t too. This is a key point for further exploration in this dissertation, and something that will be returned to. To be specific, as the 'trajectory' of short rate paths increase, it is expected that $\int_0^T r_s ds$ will increase, and thus $-\int_0^T r_s ds$ to decrease, and so $Z_T = e^{-\int_0^T r_s ds}$ will decrease.

Hence, in the case of the call option on the instantaneous short rate, there is an interplay between the pull upwards of the paths by the increased mean, the wider spread about that mean by the increased variance, and the effect of the discount factor. Similarly, the effect on the put option is that the increased mean will pull more paths out of the money, decreasing the put value. The increased spread or variance about this mean might place more paths in the money, and further in the money too, increasing the value of the put. The effect on the discount factor will also be path specific, and an accumulation of the trajectory of the short rate. In general, the increased mean would be expected to decrease the value of Z_t, but as mentioned, the spread of the short rates due to increased σ might increase the value of Z_t depending on the model. It is this that will be returned to when distributions of these random variables are explored.

3.3.3 On the Simulation of Y_t

Letting $Y_t = \int_0^t r_s ds$, to represent the integral of the simulated short rate path shows that Y_t is clearly dependent on the path of the short rate over the period $[0, t]$. An approach in simulating the Vasicek short rate is to jointly simulate r_t and Y_t, avoiding the need for a discretisation method and the introduction of another source of error. This is done through finding an analytic expression for the correlation between r_t and Y_t, allowing joint simulation. This approach to the Vasicek model is taken in Glasserman (2013).

In the Hull-White model and in the CIR models, the analytic approach that is taken in the Vasicek case, is more cumbersome and closed form expressions for the correlation not found. Thus an approximation is used – a trapezoidal quadrature scheme is then used to calculate Y_t, which is an improvement over simple quadrature. Naturally, this introduces a discretisation error, which if necessary to quantify or observe, can be deduced by comparing the mean of the Y_t simulated with that of the theoretical mean. For the theoretical distribution of Y_t in the Hull-White model, Brigo and Mercurio (2007) can be consulted.

To be clear, the trapezoidal quadrature approximate of Y_t is

$$Y_{t_i} \approx \sum_{j=1}^{i} (r_{t_{j-1}} + r_{t_j}) \frac{\Delta t}{2}.$$

Now that the random variable Y_t has been defined and an approximation scheme for it has been stipulated, the summation above can be used to see that the intuition of Y_t as the sum of the short rate values over the path taken is useful in thinking of the effect of changing σ. In the case of the Vasicek model, since σ does not change the mean of r_t, then the mean of Y_t will remain unchanged. Naturally, increased variance of r_t will also increase that of Y_t. For the Hull-White model, as the mean, and indeed the entire normal distribution of r_t shifts towards the right $\forall t \in [0, t]$, Y_t being the sum of these thus shifts rightwards, and further rightwards for higher values of σ.

3.3.4 Jensen's Inequality and Z_t

In defining the discount factor, as $Z_t = e^{-Y_t}$ for each path, an important relationship is the following:

$$\mathbb{E}[Z_t] = \mathbb{E}[e^{-Y_t}] = \mathbb{E}[e^{-\int_0^t r_s ds}] = P(0, t).$$

That is, the mean of Z_t gives the price of a ZCB maturity at t, $P(0, t)$. This mean could be estimated using a Monte-Carlo estimate (MCE), essentially an average over simulated realised paths of the short rate.

Of particular relevance to the consideration of changing σ and the effects on Y_t and subsequently on Z_t, is Jensen's inequality.

Definition 3.1. Jensen's Inequality states that if g is a convex function, then in a probabilistic sense:

$$g(\mathbb{E}[X]) \leq \mathbb{E}[g(X)],$$

where X is a random variable and $\mathbb{E}[X]$ exists.

In the specific case, where $g(x) = e^{-x}$. Then,

$$e^{-\mathbb{E}[Y_t]} \leq \mathbb{E}[Z_t] = \mathbb{E}[e^{-Y_t}] = P(0,t).$$

It is the convex relationship between Y_t and Z_t that leads to the observation that as σ increases, if $\mathbb{E}[Y_t]$ increases, then $e^{-\mathbb{E}[Y_t]}$ decreases but in such a way that this decrease is always more than the resulting effect on $\mathbb{E}[Z_t] = P(0,t)$. It is this relationship that will now be explored in the simulation and comparison of the factors mentioned here between the various models.

3.3.5 A Redefined Motivation

The observation mentioned before, that in considering a theoretical put option on the short rate and in increasing σ, firstly, there is an increase in mean and an increase in average trajectory of paths, pulling more paths upwards and out of the money. Secondly, there is also an increase in variance of these paths, placing more paths further in the money and out the money. These two effects influence the payoff. As has just been explored, another variable impacted is the discount factor, changing with σ too.

The effects deduced in considering a theoretical short rate option, can then be applied to options on bonds. Since all the models considered here are ATSM, there is the convenient inverse relationship between the short rate, r_t, and the bond price $P(t,T)$ as seen in 1.1. With this inverse relationship, as rates increase, bond prices fall. Thus, increasing rates by increasing the σ would lower bond prices. Considering an option on a bond, lower bond prices as a result of increasing σ will then decrease the value of a call option on a bond – as this would pull more bond prices out of the money. It is the resulting effects under the Vasicek model and the Hull-White model that are first explored, and later the CIR and CIR++ models.

Chapter 4

Vasicek and Hull-White: Simulation, and Pricing

The previous chapter explored the need and the method used for calibration to market data. The results of the calibration procedure indicated that the Hull-White model was capable of matching the market price of caps. It is the parameters of this calibration that are used in the Hull-White model to compare distributions of the variables mentioned at the end of the last chapter, namely that of r_t, Y_t and Z_t. The Vasicek model, if calibrated, would need to be calibrated to the bond prices or the initial term structure, and not that of the caps priced. Consequently, this difference in calibration might suggest a lack of comparison between the two. In light of this, the approach used in the simulation to follow is to leave the Vasicek model uncalibrated to the market data, and instead, to pass down the calibrated Hull-White model parameters. This approach is intended to highlight the distributional effects of changing σ in a calibrated model, and to compare with the effects in an uncalibrated model for those same values of σ.

4.1 Simulation of Short Rates

In simulating short rates, it is desirous to have various checks to validate the simulation and the coding behind the simulation. In the Vasicek and the Hull-White models paths may be simulated with an exact simulation and an Euler scheme mentioned in Chapter 3, providing said check. Furthermore, comparison with theoretical means and variances of the short rate, of bond prices, and finally with closed-form, analytic bond option prices are all methods used to validate the simulation. A few of these results will be used in this section, not only to validate the approach but also to provide a numerical perspective that serves in comparing the Vasicek and Hull-White models in the face of changing σ.

The first point of comparison is that of the prices of ZCBs. Recalling that $P(0,t) =$

$\mathbb{E}[Z_t] = \mathbb{E}[e^{-Y_t}]$, finding a Monte-Carlo estimate of the bond price at some time point t produces the price of a t-ZCB. In the Vasicek model, an endogenous model which is not calibrated to the initial yield curve, one would expect the Vasicek model to not match the market $P(0,t)$. In fact, even if the Vasicek model was calibrated to the yield curve initially, and could recover the bond prices for some value of σ, then changing σ would 'uncalibrate' it and produce different initial bond prices. That insight points to the advantages of exogenous models; if the term structure is an initial input, then changing the value of 'free' parameters has no impact on the recovery of initial bond prices, but would impact the volatility structure that it was calibrated to. This may be desirous, if considering a change in σ as a 'stress test' under which prices of caps are higher than that traded in the market. Consequently, in the Hull-White model, it is expected that the initial yield curve will be matched exactly, recovering the observed ZCB when σ is changed. This observation is explored numerically in the first table below:

	σ	2σ	5σ
	0.006656	0.013312	0.033280
Vasicek	0.997296	1.000153	1.004301
Hull-White	0.981385	0.981558	0.981228

Tab. 4.1: MCE of $P(0,t)$, where $t = 5$.

The time horizon is five years, where the bootstrapped discount factor, or equivalently, the market price of a ZCB is 0.979158519. With some error, perhaps resulting from discretisation and the calculation of Y_t as well as a potential Monte-Carlo sample size error – a sample of $n = 100000$ simulations has been used for all mean estimation and $n = 50000$ for the latter histogram plotting – it can be seen that the Hull-White ZCB price is unchanged, while the Vasicek ZCB price increases with increasing σ.

The reasons as to why this occurs, and linking the underlying intuition to the fact that the Hull-White model is calibrated, is now explored at a later time step, $t = 10$. The simple motivation behind using the last time step, $t = 10$, is that when increasing the σ parameter of the diffusion coefficient of an SDE, there is a more pronounced effect on the shifted mean of the Hull-White model at later time steps – this allows for differences in the distribution of the variables considered to be more easily seen in the corresponding histograms. Furthermore, in the choice of k, σ from calibration, the relatively low value of k, results in the effects of changing σ being more pronounced at later time steps. If k was higher, then there would be more time for the short rate to pull back towards θ in the Vasicek case, and the

deterministic value of r_t in the Hull-White that matches the initial yield curve. The effect of k on these models are explored in the limit in Chapter 2.

The following two tables show the estimation of this ZCB price, and the factors contributing to it, with changing σ under the two models. At time $t = 10$, the bootstrapped discount factor, or ZCB price, is 0.898626737.

	σ	2σ	5σ
	0.006656	0.013312	0.033280
$\mathbb{E}[r_t]$	0.001459	0.001429	0.001563
$\mathbb{E}[Y_t]$	0.009811	0.009319	0.009983
$\exp(-\mathbb{E}[Y_t])$	0.990237	0.990725	0.990066
$\mathbb{E}[Z_t] = \mathbb{E}[\exp(-Y_t)]$	0.997087	1.018461	1.176017

Tab. 4.2: Vasicek MCE, where $t = 10$.

	σ	2σ	5σ
	0.006656	0.013312	0.033280
$\mathbb{E}[r_t]$	0.022411	0.028424	0.070863
$\mathbb{E}[Y_t]$	0.109072	0.129235	0.274478
$\exp(-\mathbb{E}[Y_t])$	0.896666	0.878768	0.759969
$\mathbb{E}[Z_t] = \mathbb{E}[\exp(-Y_t)]$	0.90284	0.903254	0.901988

Tab. 4.3: Hull-White MCE, where $t = 10$.

A particularly useful comparison is to be found between the two tables above – one that centers on Jensen's inequality (3.1). In the first table, the Vasick MCEs indicate that when increasing σ, the mean of r_t is roughly unchanged, given some estimation error. As is to be seen in Figure 4.1, there is an increased variance of r_t. Both the unchanged mean and increased variance carry over to Y_t, since Y_t can simply be thought of as the sum of all r_t and an accumulation of this increased variance over the path – this is seen in Figure 4.3. The unchanged value of $\mathbb{E}[Y_t]$, when passed through the convex function, shows that $e^{-\mathbb{E}[Y_t]}$ will remain unchanged. However, in applying Jensen's inequality, $e^{-\mathbb{E}[Y_t]} \leq \mathbb{E}[Z_t]$ must hold. And as can be seen for the Vasicek model, the increase in σ affects the right hand side (RHS) of that inequality, increasing the value of $\mathbb{E}[Z_t] = P(0,t)$. The effects of changing σ on the distribution of r_t, Y_t and Z_t can correspondingly be found in the figures to follow.

Comparatively, in the Hull-White model, being calibrated to the term structure, means that the recovery of, or the matching of, the ZCB prices should be preserved under changing σ. This is seen in Table 4.3. Proceeding in a similar manner of

comparison, it can be seen that the mean of r_t increases with changing σ along with increased variance, seen in Figure 4.2. Similarly, this additive effect, is present in the increased mean of Y_t, seen in Figure 4.4. The increased expectation of Y_t, decreases the value of $e^{-\mathbb{E}[Y_t]}$. Again, applying Jensen's inequality, the convex relationship between the two infers that $e^{-\mathbb{E}[Y_t]} \leq \mathbb{E}[Z_t]$ must hold. Now, it is observed that the left hand side (LHS) of the inequality decreases, with the RHS remaining constant, leaving $\mathbb{E}[Z_t]$ to match the value of $P(0,t)$ bootstrapped.

4.2 Histograms of Distributions - changing σ

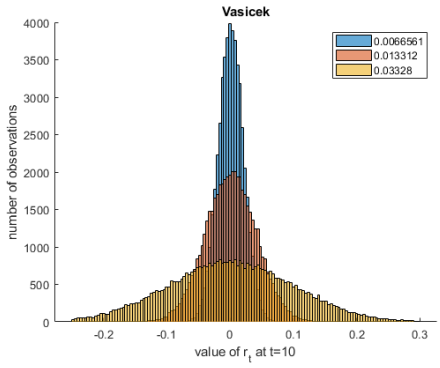

Fig. 4.1: Vasicek: Distribution of r_t

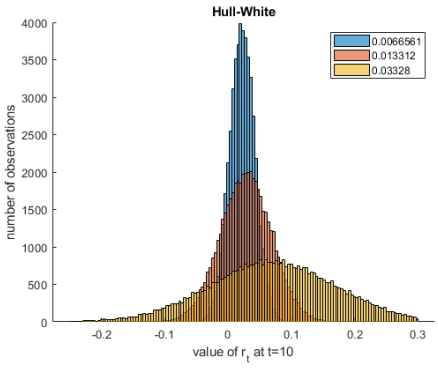

Fig. 4.2: Hull-White: Distribution of r_t

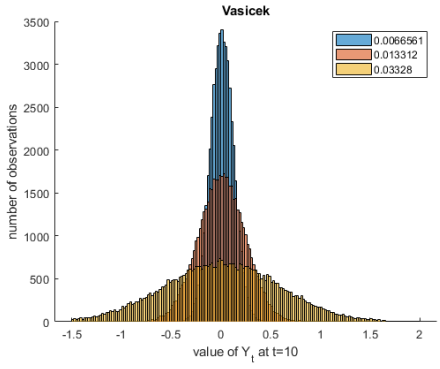

Fig. 4.3: Vasicek: Distribution of Y_t

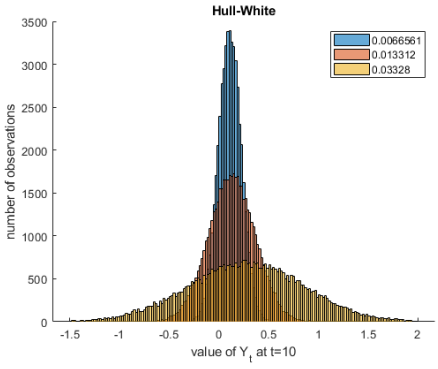

Fig. 4.4: Hull-White: Distribution of Y_t

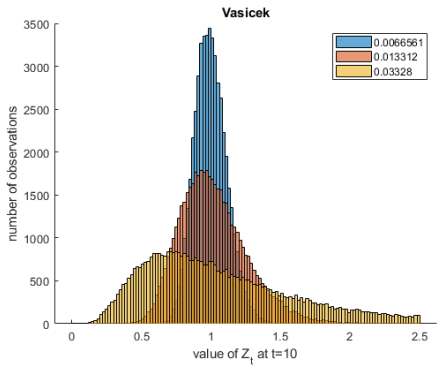

Fig. 4.5: Vasicek: Distribution of Z_t

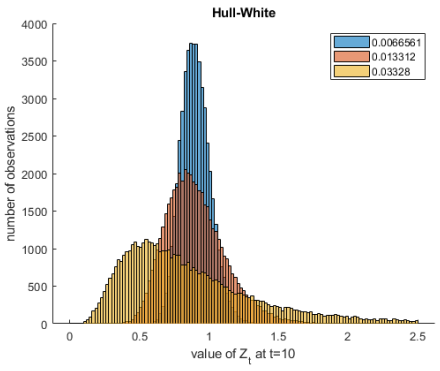

Fig. 4.6: Hull-White: Distribution of Z_t

Regarding the above histograms, of particular interest is the impact on Z_t. Since both r_t and Y_t are normally distributed under both Vasicek and Hull-White models, Z_t is lognormally distributed. As $Z_t = e^{-Y_t}$, all positive Y_t will push Z_t toward 0, this is more prevalent in the Hull-White model as the increased mean with changing σ increases r_t and avoids the negative short rates seen in the Vasick model, thus placing more Z_t closer to 0. In the Vasicek model, the greater possibility of negative short rates, when accumulated into Y_t, results in realisations of Z_t that are positive, and indeed, may be greater than 1. The effect of these negative short rates can be seen in Figure 4.5 compared to Figure 4.6. To be succinct, the increase in σ in the Vasicek model is an increase in the skewness of the lognormal distribution of Z_t. This increase in skewness shifts the mode of the distribution leftwards, and the mean rightwards, increasing $P(0,t)$. In the Hull-White model, the same alteration of the skewness occurs, however, the increase in the mean of Y_t is such that the mean of Z_t shifts leftwards by such an amount as a result of calibration, that the increase by σ results in the $\mathbb{E}[Z_t] = P(0,t)$, remaining unchanged.

4.3 Hull-White: 'Options' on Short Rates

Recalling the previous derivation of the 'theoretical' short rate options, on the value of r_T at some maturity time T, a present value at $t = 0$ is now calculated at a strike rate of $K = 0.01411$. Monte-Carlo estimates are computed, the results of which are seen in the table below. Since there is an interplay both between the payoff and the discount factor for each path, the option prices at $t = 0$ were computed for two different maturities $T = 5$ and $T = 10$. As can be seen in the table to follow, option prices for calls and puts increase with increasing σ, not displaying a decreasing option price as was theorised when looking only at the effect of σ on the payoff.

	σ	2σ	5σ
	0.006656	0.013312	0.033280
$c_0(5)$	0.003747	0.00926	0.026268
$p_0(5)$	0.008163	0.013794	0.03054
$c_0(10)$	0.010413	0.017491	0.039072
$p_0(10)$	0.00474	0.011847	0.033139

Tab. 4.4: Hull-White MCE of options on instantaneous short rate

4.4 Hull-White: Options on Bonds

In the Hull-White model, from the distributional properties of Y_t, and using the MGF of the normal distribution, the closed-form expression for the value of an S-Bond at time t is

$$\begin{aligned} P(t,S) &= \mathbb{E}[e^{-\int_t^S r_u du}|\mathcal{F}_t], \\ &= \frac{P^M(0,S)}{P^M(0,t)} e^{f^M(0,t)B(t,S) - \frac{\sigma^2}{4k}B^2(t,S)(1-e^{-2kt}) - B(t,S)r_t}. \end{aligned}$$

where $B(t,T) = \frac{1}{k}(1-e^{-k(T-t)})$ (Brigo and Mercurio, 2007).

With this expression, a Monte-Carlo estimation can also be computed for options on an S-Bond, where $S = 10$, the last maturity date of the given data set. The above expression shows that the bond price is indeed affine and is only dependent on the short rate, r_t at some time T. Thus, in order to price this bond, a similar deduction can be made about the effect of σ as was on made on r_t, except now that $P(t,T)$ is such that σ increases the mean of r_t, which has a negative coefficient in the exponent of the expression that relates $P(t,T)$ and r_t. Thus, with an increase in the mean of r_t, one would expect a decrease in the mean of the bond prices at time T, before considering the convexity effect and Jensen's inequality.

To derive bond option prices, one can follow the steps taken in Brigo and Mercurio (2007) and find the distribution of r_t under the T-forward measure $\mathbb{Q}^T$. It is from this that closed-form European call option prices on an S-bond are found, with the option maturity at time T. At some strike K, the call price at t is:

$$c(t) = P(t,S)\Phi(h) - KP(t,T)\Phi(h-\sigma_p).$$

The price of a put on the same bond, with strike and maturity the same as the above,

$$p(t) = KP(t,T)\Phi(-h+\sigma_p) - P(t,S)\Phi(-h),$$

where

$$\sigma_p = \sigma\sqrt{\frac{1-e^{-2a(T-t)}}{2a}}B(T,S),$$
$$h = \frac{1}{\sigma_p}\ln\frac{P^M(t,S)}{P^M(t,T)K} + \frac{\sigma_p}{2}.$$

The closed-form, analytic formula for these bond option prices provide a useful check to test the simulation of the short rate and subsequent pricing on, in comparing the Monte-Carlo price to that of the analytic price generated from the same parameters – this follows in the table below.

	σ	2σ	5σ	7σ
	0.006656	0.013312	0.033280	0.046593
p_0	0.070714	0.089094	0.157837	0.204903
c_0^{MCE}	0.070670	0.089540	0.156942	0.206243
p_0	0.004372	0.022752	0.091495	0.138561
p_0^{MCE}	0.004394	0.022712	0.091687	0.138901

Tab. 4.5: Hull-White analytic and MCE of options on bonds

Here the strike price was $K = 0.85$, and a sample of $n = 100000$ was used for the MCE. The MCE can be seen to be close to that of the analytic prices. Importantly, with increasing σ, prices of options on bonds increase. An additional bump of 7σ was included to show that option prices increase at this higher value of σ and to make the table more comparable with the analysis that will be carried out on the CIR++ model in the following section.

Chapter 5

CIR and CIR++: Simulation, and Pricing

An observation of CIR++ short rates is necessary to complete a comparison of the ATSMs that were introduced in the beginning of this dissertation. Additionally, the differences in extending the CIR model to the CIR++ by means of a deterministic shift, in contrast to that of the Vasicek to Hull-White, will also emerge via this comparison. Lastly, an attempt is made to find a reason as to why the CIR++ model is unable to price the market caps in the current data set, while the Hull-White model has been able to.

5.1 Choice of Parameters

A primary consideration however, is the choice of the parameters for the CIR model, as the CIR++ calibration seen in Chapter 3 was unsuccessful. The results showed a calibration attempt where, when the Feller condition was enforced, cap prices failed to match that of the market. In fact, even when the Feller condition was relaxed, despite the cap price being closer to that of the market, they were still not high enough – perhaps indicating that an upper bound in cap prices was reached.

The failure to calibrate and to find appropriate values for k and σ presents a problem; reasonable parameters are needed to simulate the CIR and CIR++ short rates. A potential solution is the use of the parameter values for the Hull-White model, where σ is now scaled up to adjust for the change in the SDE of the CIR model. The justification here is sound; as the CIR++ model has failed to calibrate to the market data, a choice of reasonable parameters are needed. With these parameters, observations can be made of the effect of changing σ.

Thus, comparing the SDE of the Hull-White (1.4), to that of the CIR++ (1.7), one can choose a σ such that the magnitude of the diffusion coefficient is the same in both models. That is, if comparing the diffusion coefficients at the initial time step

of $t = 0$, then

$$\sigma_{\text{HW}} = \sigma_{\text{CIR++}}\sqrt{r_0},$$

and if $r_0 = f(0,0) = 0.0004799$, and $\sigma_{\text{HW}} = 0.006656$, then $\sigma_{\text{CIR++}} = 0.303825$.

However the model would merely start at this value, and in fact the magnitude is not reflective of the evolution of the process. Rather, it is perhaps more appropriate to consider the mean-reverting rate that r_t will drift to over the long term, θ. Then, in this case

$$\sigma_{\text{HW}} = \sigma_{\text{CIR++}}\sqrt{\theta},$$

and if $\theta = 0.01068875$ and $\sigma_{\text{HW}} = 0.006656$, then $\sigma_{\text{CIR++}} = 0.06438055$. The second, smaller σ has been chosen as the starting value from which to bump σ upwards.

Additionally, in the plots and calculations to follow, two different sets of parameters are chosen to simulate the random variables needed. Firstly, $\alpha_1 = (k_1, \theta)$ where $k_1 = 0.0095704$, the rate of mean reversion from the calibrated Hull-White model. Secondly $\alpha_2 = (k_2, \theta)$ where $k_2 = 0.358734$ – the far higher rate of mean reversion that the CIR++ model without the Feller condition found to be closest to matching the market cap prices. The reason for simulating under different k will be best explained by the plots to follow, but primarily it is in the attempt to explain why there is an upper limit to the cap prices that the CIR++ model is able to calibrate to, and as to why calibration to the market data in Chapter 3 failed.

In the first set of plots to follow, for k_1, the narrow bounds on r_t and all subsequent variables, result in plots the scaling of which isn't ideal for display here. However, the first set of plots do not exist in isolation, and the reason for maintaining the current, slightly erroneous scaling, is to provide an easier comparison with the histograms under k_2.

5.2 Histograms of Distributions - k_1, changing σ

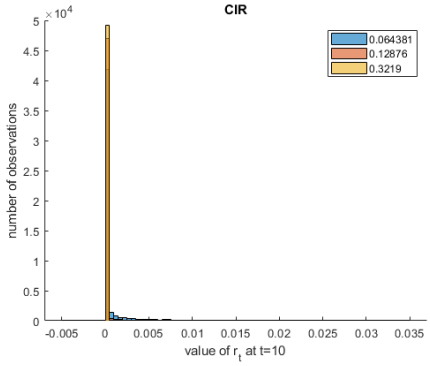

Fig. 5.1: CIR: Distribution r_t, k_1

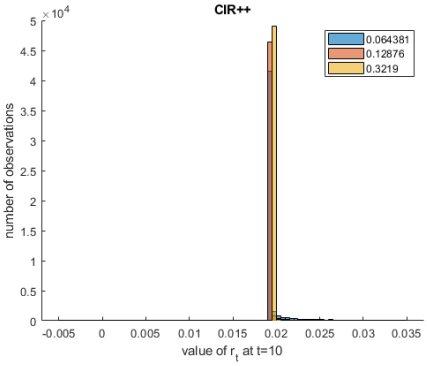

Fig. 5.2: CIR++: Distribution r_t, k_1

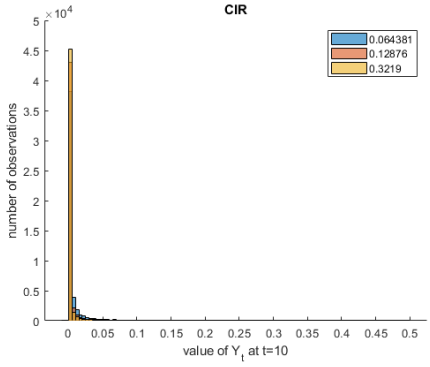

Fig. 5.3: CIR: Distribution Y_t, k_1

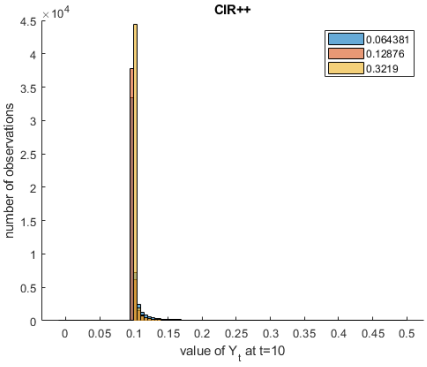

Fig. 5.4: CIR++: Distribution Y_t, k_1

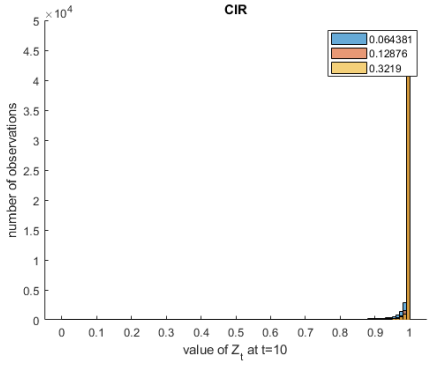

Fig. 5.5: CIR: Distribution Z_t, k_1

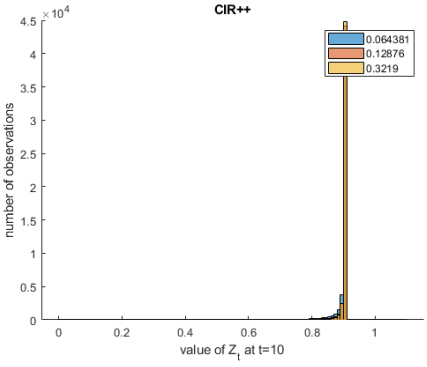

Fig. 5.6: CIR++: Distribution Z_t, k_1

5.3 Histograms of Distributions - k_2, changing σ

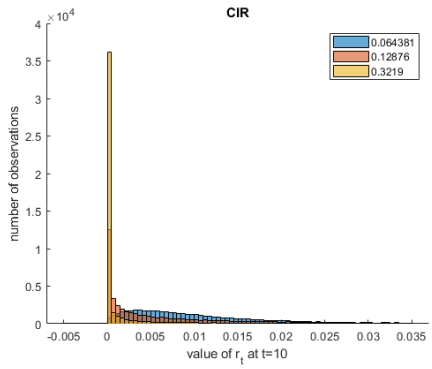

Fig. 5.7: CIR: Distribution r_t, k_2

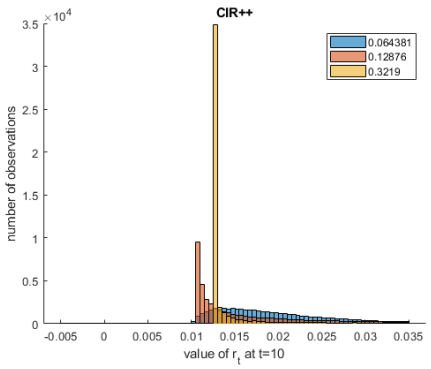

Fig. 5.8: CIR++: Distribution of r_t, k_2

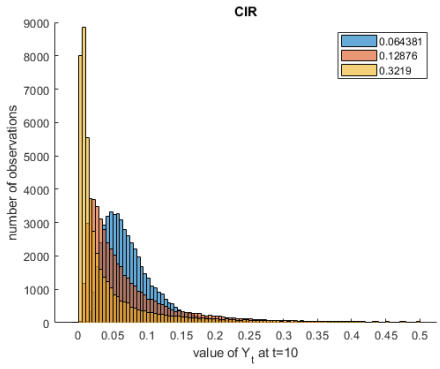

Fig. 5.9: CIR: Distribution of Y_t, k_2

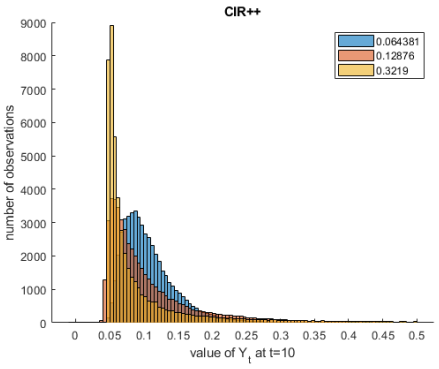

Fig. 5.10: CIR++: Distribution of Y_t, k_2

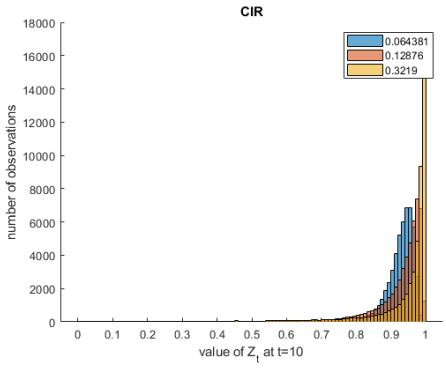

Fig. 5.11: CIR: Distribution of Z_t, k_2

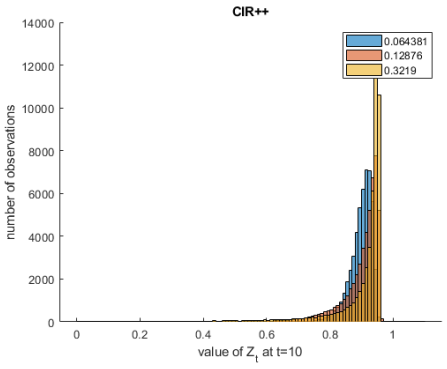

Fig. 5.12: CIR++: Distribution of Z_t, k_2

5.4 Simulation Comparison

A similar comparison may now be carried out on the means of the relevant random variables as was carried out under the Vasicek and Hull-White models. The first noticeable feature of the histograms under k_1 and k_2, is that of the deterministic shift of r_t, in moving from CIR to CIR++. This is seen first under k_1 in Figure 5.1 moving to Figure 5.2, and then more significantly under k_2 in Figure 5.7 to 5.8.

In the histograms of the short rate, a crucial observation can be made that centers on the effect of k on the distributions of r_t. Firstly, r_t cannot be negative, consequently, there is a lower bound on r_t of 0 in the CIR model, which then becomes a lower bound of $\varphi(t)$ where $t \in [0, 10]$ when moving to the CIR++ model. Secondly, as $\theta = 0.010499253$ is low, a lower rate of mean reversion k results in more and more of r_t being around 0 in the CIR model, seen clearly in Figure 5.1, contrasted with Figure 5.7. Importantly, the value of $\mathbb{E}[r_t]$ is much higher under k_2 than it is under k_1, in line with what was expected of $\mathbb{E}[r_t]$ when analysed for the CIR model in Chapter 2. Furthermore, with low k and low θ, there is a lower mean and the short rate sits mostly around 0. With higher k, the higher rate of mean reversion works to overcome the higher variance, and shifts the distribution around θ – this is seen in Table 5.2 below and in the histograms under k_2.

k_1	CIR			CIR++		
	σ	2σ	5σ	σ	2σ	5σ
$\mathbb{E}[r_t]$	0.00145	0.00149	0.00281	0.02052	0.02085	0.02271
$\mathbb{E}[Y_t]$	0.00995	0.01117	0.02322	0.10724	0.10949	0.12449
$\exp(-\mathbb{E}[Y_t])$	0.99010	0.98889	0.97705	0.89831	0.89629	0.88295
$\mathbb{E}[Z_t]$	0.99055	0.99067	0.98923	0.89872	0.89791	0.89395

Tab. 5.1: CIR and CIR++ MCE under k_1, where $t = 10$.

k_2	CIR			CIR++		
	σ	2σ	5σ	σ	2σ	5σ
$\mathbb{E}[r_t]$	0.01043	0.01054	0.01048	0.02051	0.02099	0.02268
$\mathbb{E}[Y_t]$	0.07956	0.07942	0.08238	0.10738	0.10908	0.12162
$\exp(-\mathbb{E}[Y_t])$	0.92352	0.92366	0.92092	0.89818	0.89666	0.88549
$\mathbb{E}[Z_t]$	0.92413	0.92595	0.93285	0.89877	0.89889	0.89696

Tab. 5.2: CIR and CIR++ MCE under k_2, where $t = 10$.

When considering the CIR++ model, as mentioned, the first thing to notice is the shift in distribution as a result of φ. The table below clearly shows the effect

of φ under the different values of k. The higher $\mathbb{E}[r_t]$ in the CIR model under k_2, results in a smaller deterministic shift, φ, to the CIR values. It is this shift that allows calibration; both Table 5.1 and Table 5.2 show that $\mathbb{E}[Z_t]$ is roughly equal to $P(0,t) = 0.898626737$, showing that the CIR++ models does indeed recover the observed ZCB price for maturity $t = 10$.

The consequences of the higher k are revealing; in the CIR model the higher k ensures that the distribution is centred around θ and not 0, as it is with k_1. It is then the increased amount of non-zero rates which allows the higher σ to affect the paths simulated – this is the striking difference seen between the histograms and the increased spread of the distributions under k_2. These distributions are then shifted to the CIR++ model, from Figure 5.7 to Figure 5.8, where the effect of φ is far more striking than under k_1, acting now on markedly different distributions.

	σ	2σ	5σ
	0.064381	0.128761	0.321903
		k_1	
$\varphi(10)$	0.01908	0.01936	0.01990
$\mathbb{E}[r_t]^{\text{CIR++}} - \mathbb{E}[r_t]^{\text{CIR}}$	0.01908	0.01936	0.01990
		k_2	
$\varphi(10)$	0.01008	0.01045	0.01221
$\mathbb{E}[r_t]^{\text{CIR++}} - \mathbb{E}[r_t]^{\text{CIR}}$	0.01008	0.01045	0.01221

Tab. 5.3: Difference between r_t indicating shift by φ

Lastly, an observation of Z_t is crucial to understanding the calibration under the CIR++ model. In the CIR model the lower bound of 0, as mentioned, provides an upper bound of 1 for Z_t, since $Z_t = \exp(-Y_t)$ and as Y_t is always positive, hence $0 \leq Z_t \leq 1$. This is most clearly seen in Figure 5.12. For the CIR++ model, the aforementioned determinsitic shift affects each instantaneous short rate across $t \in [0, 10]$ and the accumulation of those shifts is then seen in Y_t. The shift likewise shifts Z_t. Since Y_t would be increased, resulting in larger positive values, this then moves the lower bound of 0 for Y_t under the CIR model. Y_t is then shifted similarly under the CIR++ model which then effects the the values of Z_t; the higher lower bound of Y_t similarly providing a lower upper bound for Z_t , which can be seen in comparing Figure 5.7 to Figure 5.8.

A similar application of Jensen's inequality due to the convex relationship between Y_t and Z_t can be done in this section too, but due to the fact that the CIR++ is a shifted CIR, the usefulness of such an explanation is perhaps of limited value in the light of the description that has preceded above. That being said, what was ob-

served in Chapter 4 still holds here. In the CIR model, it was deduced that the theoretical mean of r_t is independent of σ, which is seen in the tables above. Firstly in Table 5.1 a relatively constant $\mathbb{E}[r_t]$ is observed, where the last, higher value under 5σ might very well be due to discretisation errors since the paths simulated are so close to 0; the theoretical mean under k_1 is $0.001411...$. Secondly, under k_2, Table 5.2 shows a more constant $\mathbb{E}[r_t]$, roughly around the theoretical mean of $0.0104062...$, calculated from the theoretical expression in Chapter 2. For Jensen's Inequality (3.1) to hold, $e^{-\mathbb{E}[Y_t]} \leq \mathbb{E}[Z_t]$. In the CIR model, this results in $\mathbb{E}[Z_t]$ mostly increasing (under k_2 it is noticeable) with increasing σ as was the case in the Vasicek model. When moving to the CIR++ model, the deterministic shift increases with increasing σ, seen in Table 5.3, which then increases $\mathbb{E}[r_t]$. These effects result in the left-hand side of Jensen's inequality, $e^{-\mathbb{E}[Y_t]}$, decreasing, with the right-hand side, $\mathbb{E}[Z_t]$ remaining constant. This is similar to the effect under the Hull-White model and ensures that the CIR++ model is calibrated.

5.5 Comparison of $P(T,S)$

The histograms of Z_t hints at something important happening to the distribution of bonds under different parameter sets. Consequently, a comparison of $P(T,S)$ is of interest, with attention paid to the upper and lower bounds of $P(T,S)$. Here, $P(T,S)$ is the time $T = 5$ price of a bond maturing at time $S = 10$. This is the same maturity, T, of the same S-Bond that was calculated in Chapter 4. Here, the calculation of $P(T,S)$ can best be summarised as follows:

Under the CIR model,

$$P(t,T) = A(t,T)e^{-B(t,T)r_t},$$

affirming that the CIR model is indeed affine, and

$$A(t,T) = \left[\frac{2he^{(k+h)(T-t)/2}}{2h + (k+h)(e^{(T-t)h} - 1)}\right]^{\frac{2k\theta}{\sigma^2}},$$
$$B(t,T) = \frac{2(e^{(T-t)h} - 1)}{2h + (k+h)(e^{(T-t)h} - 1)},$$

where h is defined as in 1.7, that is $h = \sqrt{k^2 + 2\sigma^2}$.

Under the CIR++ model,

$$P(t,T) = A_{\text{CIR}}(t,T)e^{-B(t,T)r_t},$$

where

$$A_{\text{CIR}}(t,T) = \frac{P^M(0,T)A(0,t)e^{-B(0,t)x_0}}{P^M(0,t)A(0,T)e^{-B(0,T)x_0}}A(t,T)e^{B(t,T)\varphi(t)},$$

where $P^M(0,T)$ is the market price of a ZCB maturing at T from the dataset to be calibrated to (Brigo and Mercurio, 2007). Recall that $r_t = x_t + \varphi^{\text{CIR}}(t)$ as in Equation 1.7.

The formulae above show that for $P(t,T)$ the relationship between the CIR and CIR++ model is also that of a determinsitc shift – as taking the ratio of the above expressions for $P(t,T)$ would reveal. Furthermore, as $r_t \geq 0$ under both of these models, there is an upper bound on $P(t,T)$. For the CIR model, the upper bound is $A(t,T)$ and for the CIR++ model, it is $A_{\text{CIR}}(t,T)$. The upper bound on $P(T,S)$ as well as the deterministic shift that changes with σ is most clearly seen in Figure 5.16 below. Interestingly, this shows that under the CIR++ model, the upper bound for 5σ is in fact lower than for 3σ, resulting in the distribution for $P(T,S)$ lying to the left of the one of the lower σ.

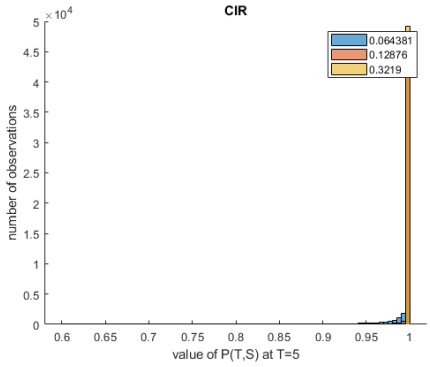

Fig. 5.13: CIR: $P(T,S)$ under k_1

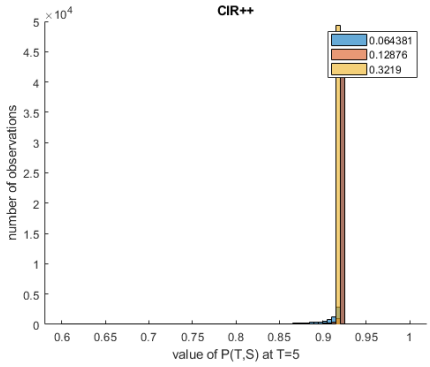

Fig. 5.14: CIR++: $P(T,S)$ under k_1

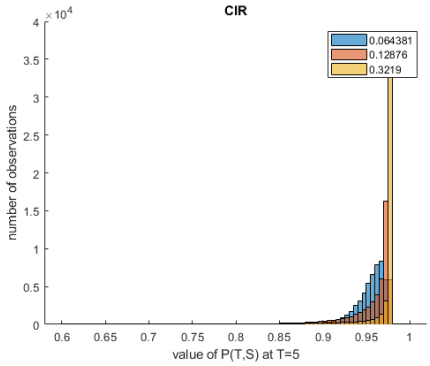

Fig. 5.15: CIR: $P(T,S)$ under k_2

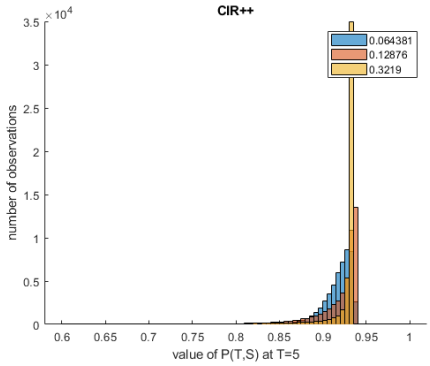

Fig. 5.16: CIR++: $P(T,S)$ under k_2

Performing a similar analysis on the MCE of the means of $P(T,S)$ and of the upper bound (which is a closed form expression and requires no estimate) now follows in the tables below. Table 5.4 shows an increasing mean with increasing σ,

that is more easily seen under the higher k_2 – this is inline with the derivation by Cox *et al.* (1985) that $P(t,T)$ is an increasing convex function of σ^2. This table also shows that the upper bound increases with σ, moving closer and closer to 1.

Under the CIR++ model, $\mathbb{E}[P(T,S)]$ is now a decreasing function of σ for this set of input rates. This is seen in Table 5.5 where it is also seen that the upper bound is a decreasing function of σ. With this information, the effect of calibration on options on bonds is now analysed.

	σ	2σ	5σ
	0.064381	0.128761	0.321903
		k_1	
$\mathbb{E}[P(T,S)]$	0.994124	0.994389	0.994330
$A(T,S)$	0.998753	0.998783	0.998942
		k_2	
$\mathbb{E}[P(T,S)]$	0.951926	0.953219	0.958271
$A(T,S)$	0.971910	0.972231	0.974165

Tab. 5.4: CIR: MCE and upper bound of $P(T,S)$

	σ	2σ	5σ
	0.064381	0.128761	0.321903
		k_1	
$\mathbb{E}[P(T,S)]$	0.917435	0.916867	0.914714
$A_{\text{CIR}}(t,T)$	0.960750	0.958434	0.948658
		k_2	
$\mathbb{E}[P(T,S)]$	0.917478	0.917360	0.915374
$A_{\text{CIR}}(t,T)$	0.938129	0.937421	0.934086

Tab. 5.5: CIR++: MCE and upper bound of $P(T,S)$

5.6 CIR++: Options on Bonds

The distributions of $P(T,S)$ indicate something striking, that there is an upper bound on $P(T,S)$ that decreases with increasing σ under this particular term structure. Call and put prices on this S-Bond are now calculated under k_1 and k_2, comparing analytic prices to MCE to provide a justification of the results observed. Another additional bump of σ has been added in this comparison that was not present in the previous histograms; this has been done to compare option prices at higher σ and postulate a reason as to why the attempted calibration in Chapter 3 returned

a value of $\sigma = 358734$.

Firstly, under k_1, Table 5.6 clearly show that call and put prices decrease at $5\sigma = 0.321903$, noting that the point at which option prices decrease could potentially be at a lower value of σ. Under k_2, Table 5.7 shows that this decrease is later, at higher values of σ – occurring at $7\sigma = 0.450664$.

Interestingly, while the bond option prices are slightly lower under k_2 at lower σ, they increase by more as one moves to 2σ. Furthermore, the fact that the prices only decrease at later σ shows that the option prices under k_2 can reach higher values at higher σ.

	σ	2σ	5σ	7σ
	0.064381	0.128761	0.321903	0.450664
c_0	0.019193	0.019671	0.018407	0.017949
c_0^{MCE}	0.019140	0.019598	0.018317	0.017824
p_0	0.001809	0.002287	0.001023	0.000565
p_0^{MCE}	0.001918	0.002910	0.001767	0.001028

Tab. 5.6: CIR++ options on bonds, k_1, $K = 0.9$

	σ	2σ	5σ	7σ
	0.064381	0.128761	0.321903	0.450664
c_0	0.019069	0.022942	0.025405	0.024008
c_0^{MCE}	0.019017	0.022934	0.025362	0.023892
p_0	0.001685	0.005558	0.008021	0.006623
p_0^{MCE}	0.001706	0.005693	0.008141	0.006848

Tab. 5.7: CIR++ options on bonds, k_2, $K = 0.9$

It is the above two observations, combined with the histograms of $P(T, S)$ that indicate that the upper bound on $P(T, S)$ that decreases with higher σ, results in decreasing bond option prices with increasing σ. It is this that would explain the failure of the CIR++ model to calibrate to the cap prices in Chapter 3, as an upper bound for a put price, results in an maximum possible caplet price, thus, an upper bound on a cap price.

Chapter 6

Conclusion

This dissertation explored two groupings of ATS models. The first is that of the Vasicek model and the exogenous extension, the Hull-White model. The second is that of the CIR model and its extension, the CIR++ model. In moving from endogenous to exogenous models, changing σ affects the mean of the short rates, r_t. With this observation, potential consequences on option prices were theorised – the focus, the potential of decreasing bond option prices when increasing σ. This would be a counter-intuitive result, as stress-tests are immediately thought to increase the value of options.

The exogenous models were calibrated to a volatility structure from Euribor market data, with the Hull-White model being able to match market caps sufficiently, while the CIR++ model failed to do so. When simulations with calibrated parameters were performed, the Hull-White model showed increasing option prices across all values of σ. In the second group, the CIR++ model showed decreasing bond option prices at very reasonable bumps of σ. Furthermore, another observation at higher k, showed decreasing option prices but at higher σ. The reason for the decrease in option prices with increasing σ stems from the combination of the underlying model, the CIR model, and the deterministic shift that increases the mean of r_t. The CIR model does not allow negative rates and this lower bound on r_t, produces an upper bound on $P(T, S)$. The deterministic shift in moving to the exogenous model, results in a higher lower bound under the CIR++ for r_t, and a lower upper bound on $P(T, S)$ that decreases with increasing σ. It is this that causes decreasing bond option prices after a certain value of σ.

The consequences of this are twofold and yet fundamentally tied. Firstly, when stress-testing the CIR++ model, consideration needs to be made for what value σ is bumped to, as was shown that depending on the parameters used, even reasonable bumps can produce lower bond option prices. Secondly, the later decreasing prices of bond options explains why there is an observed upper bound on cap prices, resulting in failed CIR++ calibration to the market volatility structure.

Bibliography

Benninga, S. and Wiener, Z. (1998). Term structure of interest rates, *Mathematica in Education and Research* **7**: 13–21.

Björk, T. (2009). *Arbitrage theory in continuous time*, Oxford University Press.

Brigo, D. and Mercurio, F. (2000). The CIR++ model and other deterministic-shift extensions of short rate models, *Proceedings of the 4th Columbia-JAFEE conference for mathematical finance and financial engineering*, pp. 563–584.

Brigo, D. and Mercurio, F. (2007). *Interest rate models-theory and practice: with smile, inflation and credit*, Springer.

Cox, J. C., Ingersoll Jr, J. E. and Ross, S. A. (1985). An intertemporal general equilibrium model of asset prices, *Econometrica: Journal of the Econometric Society* pp. 363–384.

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates, *Mathematical Finance* **6**(4): 379–406.

Filipovic, D. (2009). *Term-Structure Models. A Graduate Course.*, Springer.

Glasserman, P. (2013). *Monte Carlo methods in financial engineering*, Vol. 53, Springer Science & Business Media.

Ho, T. S. and Lee, S.-B. (1986). Term structure movements and pricing interest rate contingent claims, *Journal of Finance* **41**(5): 1011–1029.

Hull, J. and White, A. (1990). Pricing interest-rate-derivative securities, *Review of Financial Studies* **3**(4): 573–592.

Lord, R., Koekkoek, R. and Dijk, D. V. (2010). A comparison of biased simulation schemes for stochastic volatility models, *Quantitative Finance* **10**(2): 177–194.

Piazzesi, M. (2010). Affine term structure models, *Handbook of Financial Econometrics: Tools and Techniques*, Elsevier, pp. 691–766.

Vasicek, O. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**(2): 177–188.